Report from Dagstuhl Seminar 25101

Guardians of the Galaxy: Protecting Space Systems from Cyber Threats

Ali Abbasi*[1], Gregory J. Falco*[2], Daniel Fischer*[3], and Jill Slay*[4]

1 **CISPA Helmholtz Center for Information Security, DE.** abbasi@cispa.de
2 **Cornell University – Ithaca, US.** gfalco@cornell.edu
3 **ESA / ESOC – Darmstadt, DE.** daniel.fischer@esa.int
4 **University of South Australia – Mawson Lakes, AU.** jill.slay@unisa.edu.au

Abstract

This report documents the program and outcomes of Dagstuhl Seminar 25101 "Guardians of the Galaxy: Protecting Space Systems from Cyber Threats," which brought together 40 participants from 11 countries. It explains why space cybersecurity is distinct from terrestrial contexts and distills the working-group results (attack/prepare, detect, protect, respond) into a focused research-and-action roadmap for agencies, industry, and academia.

Seminar March 02–07, 2025 – https://www.dagstuhl.de/25101

2012 ACM Subject Classification Computer systems organization → Embedded and cyber-physical systems; Security and privacy → Security in hardware; Security and privacy → Systems security; Networks → Network security

Keywords and phrases Space Cybersecurity, Satellite Security, Cyber-Physical Systems, Network Security, Embedded Systems Security, System Security, Autonomous Systems Security, Post-Quantum Cryptography

Digital Object Identifier 10.4230/DagRep.15.3.1

1 Executive Summary

Ali Abbasi
Gregory J. Falco
Daniel Fischer
Jill Slay

This report synthesizes the outcomes of Dagstuhl Seminar 25101, "Guardians of the Galaxy: Protecting Space Systems from Cyber Threats," which convened 40 experts from academia, industry, and government. The seminar established a clear consensus that space cybersecurity is a qualitatively distinct discipline, not merely an extension of terrestrial challenges. The seminar focused on:

- **Defining the Foundational Challenges:** Articulating why space is different and how this affects the security domain, focusing on the ambiguity created by the harsh physical environment, the necessity of high-stakes autonomy due to extreme latency, and the uniquely asymmetric attack surface.
- **Structuring the Problem Space:** Organizing analysis and solutions around four key operational functions via dedicated working groups: ATTACK/PREPARE, DETECT, PROTECT, and RESPOND.

* Editor / Organizer

Guardians of the Galaxy: Protecting Space Systems from Cyber Threats, *Dagstuhl Reports*, Vol. 15, Issue 3, pp. 1–38
Editors: Ali Abbasi, Gregory J. Falco, Daniel Fischer, and Jill Slay
Dagstuhl Reports
Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

- **Formulating a Strategic Roadmap:** Proposing a multi-pillar plan to foster a cumulative and collaborative research ecosystem that bridges the gap between academic innovation and operational needs.

As a major result, the seminar identified the following interconnected problem areas and corresponding future research directions:

1. **The Testbed and Data Gap:** Overcoming the critical shortage of realistic research infrastructure by developing a federated ecosystem of high-fidelity testbeds. A key requirement is that these testbeds must be segment-complete, modeling the entire ground-link-space chain, and support graduated fidelity. This allows researchers to move between pure simulation, hardware-in-the-loop, and testing with unmodified firmware binaries depending on the research question. Furthermore, institutional spacecraft operators should be encouraged to share more representative data sets that can be used in research.
2. **Securing Next-Generation Communications:** Addressing the unique security needs of future space networks. This includes maturing protocols for the Solar System Internet (e.g., Delay Tolerant Networking – DTN) essential for deep space, planning the transition to Post-Quantum Cryptography (PQC), and developing resilient defenses against jamming and spoofing for high-bandwidth optical and RF links.
3. **Building Trustworthy Autonomous Systems:** Ensuring that onboard AI and autonomous systems are secure and safe. This requires developing physics-informed and resource-aware intrusion detection, designing systems to be “forensic-by-design” so that evidence of an attack survives recovery actions, and implementing verifiable and secure software update pipelines.
4. **Strengthening the System Foundation:** Mandating a “secure-by-design” philosophy anchored in hardware. This involves adopting measured boot processes, internal message-level authentication, and robustly managing the cybersecurity of the global supply chain (C-SCRM) for all components.
5. **Establishing a Collaborative Ecosystem:** Creating the necessary non-technical structures for progress. This includes developing clear governance and interoperable standards, establishing “safe-harbor” policies for vulnerability disclosure, and implementing new collaborative models, such as co-funded PhD programs, to grant researchers vital access to realistic systems and data.

2 Table of Contents

Executive Summary

Overview of Talks

3 Overview of Talks

3.1 Powering Europe's Space Ambition: Cybersecurity Challenges in Space Systems

Daniel Fischer (ESA / ESOC – Darmstadt, DE, daniel.fischer@esa.int)

The European Space Agency (ESA) is responsible for the peaceful exploitation of space on behalf of its member states. It is active in all major domains of space, from launchers, human spaceflight, earth observation, and GNSS, to science and communication. Many of the systems developed by ESA, either directly on behalf of its member states, or on behalf of the European Commission (e.g., Galileo, Copernicus, IRIS2), represent critical infrastructure upon which society depends on a daily basis.

Cybersecurity has thus grown to be a major challenge in the development of ESA programs and assets, in particular in today's changing geopolitical landscape. In response to these challenges, ESA has made cybersecurity one of its three main technology priorities in addition to quantum and AI.

The quick development and maturation of space security technologies, together with the European space industry and academia, is fundamental. For this purpose, ESA seeks to connect closer with these entities and exploit synergies. ESA seeks to supply the academic ecosystem with relevant space use cases while benefitting from the resulting research to speed up technology spin-in. Likewise, industry is a valuable partner in picking up the higher technology readiness level (TRL) developments in cybersecurity and creating a diverse ecosystem for operational cybersecure space system assets and components.

3.2 Cybersecurity Challenges in Space Systems: Notable Challenges and Research Areas

Marcus Wallum (ESA / ESOC – Darmstadt, DE, marcus.wallum@esa.int)

The talk presented an overview of current challenges and potential future research topics. Topics included :

- Digital security engineering, alignment with Model-based System Engineering and formal reference architectures
- Zero trust architectures for space systems
- Post-Quantum Cryptography, its impact on space communications and need for cryptographic agility
- Tailored space system security monitoring and testing solutions including fuzzing of space communication protocols
- Securing legacy systems
- Anomaly detection and responsive resilient self-healing architectures
- Securing the supply chain
- Evolution of avionics security architectures and their secure operation, including on-board IDS/IPS, TEE, remote attestation
- Leveraging AI for security and ensuring secure use of AI
- Applied confidential computing and homomorphic encryption for secure distributed dataset processing
- Proliferation of standards, regulations and certification scheme

3.3 Space System Security and the Space Environment

Knut Eckstein (ESA / ESTEC – Noordwijk, NL, knut.eckstein@esa.int)

The talk aimed at initiating fruitful discussions between academics and practitioners by positing which aspects of space systems security engineering are the most challenging or the most interesting from an academic Research and Development perspective. It started by noting that spacecraft, compared to other mobile network nodes, have neither the least powerful CPUs, nor the smallest amounts of memory, nor the least predictable communication network topologies, nor the longest periods of communication outages. What is special about spacecraft is that their wireless links are highly asymmetric in nature and are absolutely essential, in absence of any wired links that can be established in drones or aircraft during maintenance phases. Also, spacecraft are fairly unique in their focus of safety and availability over long periods of time without "return to base" i.e. any security mechanism design has to satisfy very stringent safety requirements.

3.4 Down to Earth: Cyber Security Operations

Markus Rückert (ESA / ESOC – Darmstadt, DE, markus.rueckert@esa.int)

Following the NIST Cyber Security Framework (CSF), the talk summarized ESA's approach to PROTECT, DETECT, and RESPOND at a conceptual level and in order to protect ESA's Operations, Investments, and Brand Value. The talk created awareness of sector-specific cyber security challenges with the aim of stimulating the ideation for seminar topics.

The talk illustrated the nature and diversity of the assets (infrastructure, services, information) that require protection.

Furthermore, the talk highlighted a series of key challenges in the context of cyber-physical systems as opposed to traditional IT.

The widespread use of shared ground infrastructures, due to cost benefits, exposes a wide attack surface, making it harder to protect from cyber threats. Complex and highly specialized supply chains present unique challenges when it comes to effective identification and management of weaknesses and vulnerabilities, as well as when it comes to the identification of threats and countermeasures. Similarly, the presence of dual-use technologies may limit information sharing among the parties involved. In general, system complexity and interoperability constraints often slow the adoption of new technologies, including improved security controls and protection practices.

In general, the resulting inertia and obstacles affect the evolution of PROTECT, DETECT, and RESPOND. There is a general call for research, developmentand innovation to take these factors into aaccount.

3.5 Hack The Planet and Beyond: Security Challenges of the Solar System Internet (SSI)

Lars Baumgärtner (ESA / ESOC – Darmstadt, DE, lars.baumgaertner@esa.int)

The SSI is built upon new protocols, technologies, and mechanisms, particularly the concept of ‘store-carry-and-forward’ (SCF) for Delay-Tolerant Networking (DTN). While this approach addresses the fluctuating connectivity and high delays inherent in interplanetary communication, it also creates the need for novel security solutions and prevents the use of existing security measures. Several key areas present major challenges:

- Delay-tolerant key management
- SSI Threat Modelling
- Delay-tolerant networking (DTN) Anomaly Detection
- Security of inter-planetary multicast
- Scalable network testbeds for SSI

3.6 NASA Mission Resilience & Protection Approach – Including space Cybersecurity

Kevin Gilbert (NASA Goddard Space Flight Center – Greenbelt, US, kevin.w.gilbert@nasa.gov)

This talk provides an overview of NASA STD-1006A (NASA’s space protection requirements), then gives an overview of the NASA protection planning process (which includes Candidate Protection Strategies related to space mission cybersecurity), and will conclude with a snapshot of where we think development is needed to find protection solutions for civil space missions.

3.7 Security Units for Satellite Communication | Challenges

Arne Grenzebach (OHB System – Bremen, DE, arne.grenzebach@ohb.de)

This talk presents the current challenges of developing security units for satellite communication. This is based on industrial experience within a satellite manufacturing company, namely OHB.

3.8 Space Attack Research and Tactic Analysis (SPARTA)

Brandon Bailey (The Aerospace Corp. – Los Angeles, US, brandon.bailey@aero.org)

This talk presents an overview of the Tactic Technique, & Procedure (TTP) framework called SPARTA. We describe how it can be used to document attacks on spacecraft along with countermeasures to mitigate or prevent the attacks. The goal was education and awareness of the tool & present future capabilities. SPARTA was the first of its kind repository of knowledge on how to attack or defend spacecraft.

3.9 Migrating Legacy Ground Stations to Cloud-based Zero-trust Stations

Mattias Wallén (Swedish Space Corporation – Solna, SE, mattias.wallen@sscspace.com)

This talk presents current threats and vulnerabilities to satellite ground stations. The talk was focused on the need to move to cloud-based ground stations and reduce risk by using DevSecOps, loosely coupled systems, Zero trust architectures, policy as code, infrastructure as code, and compliance as code. Compared to the physical industry, computer science, and IT – the space industry is still in a "Steam Power" state and moving towards assembly line and automation.

3.10 New Space = Secure Space?

Steven Arzt (Fraunhofer SIT – Darmstadt, DE, steven.arzt@sit.fraunhofer.de)

The space industry is changing with the "New Space" activities, new technologies and new business models challenge traditional risk models and security measures. As part of the expert group on space security by BSI (Germany's Federal Office for Information Security), we look into this evolving landscape. Further, governmental missions on new topics such as QKD, space debris, and AI-driven security analysis require us to change existing solutions and insights. What would a world in which anyone can launch a satellite a rent a ground station look like security-wise?

Lastly, we need to bring more bright minds into the intersection of space and cybersecurity. Hacking contests and CTFs can bridge the path into the field and reduce the barrier of entry.

3.11 A Joint Effort: Stakeholder Cooperation for Better Cybersecurity in Space

Florian Göhler (BSI – Bonn, DE, florian.goehler@bsi.bund.de)

Cybersecurity needs to be an integrated part of every space mission, and security aspects should be considered throughout all phases of a project. However, there was a lack of regulation and security standards that address cyber threats in space. To overcome this issue, the German Federal Office for Information Security founded an expert group for cybersecurity in space that invites experts from governmental institutions, industry, and academia to work together on standardization and regulation. In this joint effort, the expert group developed multiple documents that aim to mitigate cyber threats on space and ground segments. Furthermore, the expert group aims to identify emerging new technologies and regulations that may impact cybersecurity in space. These efforts also take international developments into account. This talk will give an overview of the activities of the group and its security documents.

3.12 Merge/Space: A Security Testbed for Satellite Systems

Stephen Schwab (USC/ISI – Arlington, US, schwab@isi.edu)

Merge/Space (M/S) is a testbed designed to simulate multiple-agent security scenarios in satellite networks. By combining orbital data generated by a simulator such as STK with a synchronized set of images, M/S can accurately simulate bandwidth and connectivity constraints between ground stations and vehicles, enabling analyses of DoS attacks, scanning, malware infiltration, and other analyses. We discuss the development of the testbed, and the sample datasets included for release, and demonstrate the impact of various simulations.

3.13 HoneySat: A Network-based Satellite Honeypot Framework

Efrén López-Morales (Texas A&M University – Corpus Christi, US, elopezmorales@islander.tamucc.edu)

Joint work of Efrén López-Morales, Ulysse Planta, Ali Abbasi

Satellites are the backbone of several mission-critical services such as GPS that enable our modern society to function. For many years, satellites were assumed to be secure because of their indecipherable architectures and the reliance on security by obscurity. However, technological advancements have made these assumptions obsolete, paving the way for potential attacks, and sparking a renewed interest in satellite security. Unfortunately, to this day, there is no efficient way to collect data on adversarial techniques for satellites, which severely hurts the generation of security intelligence. In this paper, we present HoneySat, the first high-interaction satellite honeypot framework, which is fully capable of convincingly simulating a real-world CubeSat, a type of Small Satellite (SmallSat) widely used in practice. To provide evidence of the effectiveness of HoneySat, we surveyed experienced SmallSat operators currently in charge of active in-orbit satellite missions. Results revealed that the majority of satellite operators (71.4%) agreed that HoneySat provides realistic and engaging simulations of CubeSat missions. Further experimental evaluations also showed that HoneySat provides adversaries with extensive interaction opportunities by supporting the majority of adversarial techniques (86.8%) and tactics (100%) that target satellites. Additionally, we also obtained a series of real interactions from actual adversaries by deploying HoneySat on the internet over the span of several months, confirming that HoneySat can operate covertly and efficiently while collecting highly valuable interaction data.

3.14 Securing the Satellite Software Stack

Samuel Jero (MIT Lincoln Laboratory – Lexington, US, samuel.jero@ll.mit.edu)

Satellites and the services enabled by them play an increasingly important in our modern life. To support these services, satellite software is becoming increasingly complex and connected. As a result, concerns about its security are becoming prevalent. While the focus of security has historically been encrypting communication links, we argue that further consideration of the security of satellites is necessary. This talk characterizes the cyber threats to satellites, surveys the unique challenges for satellite software, and presents a vision for future research in this area.

3.15 Developing accessible test beds and data sets

Jill Slay (University of South Australia – Mawson Lakes, AU, jill.slay@unisa.edu.au)

To expand and extend the growing area of satellite cybersecurity to larger and more diverse cohorts of cross-disciplinary researchers internationally, we need appropriate datasets and test beds where developed protection solutions can be studied. The emerging challenge is to standardize such research infrastructure to begin to answer wicked space cyber research questions so as to protect humans and their space missions.

3.16 On the Security of Non-Terrestrial Networks

Gunes Karabulut Kurt (Polytechnique Montréal, CA, gunes.kurt@polymtl.ca)

6G networks are expected to be a combination of the terrestrial network and the non-terrestrial network (NTN). Elements of NTN will be base stations with 3D mobility, such as low Earth orbit (LEO) satellites, unmanned autonomous vehicles (UAVs), and high altitude platform station (HAPS) systems. The presence of such NTN elements introduces new features in terms of coverage, computation, localization, and sensing. However, their presence also makes 6G networks vulnerable to new security threats, especially in the physical layer (PHY). After detailing the NTN evolution, this talk focuses on two different threats. The first threat type emerges from the communication attacks that are expected to increase with the presence of wireless backhaul connectivity. The second threat type is on the localization systems, especially for NTN elements, as the location information of a LEO satellite, a UAV, or a HAPS is an essential network characteristic that will affect the overall network performance. The talk will conclude with the importance of physical layer security for NTNs, an overview of the open issues, and future research directions.

4 Open Problems

Space-security research is hindered by a tooling gap: there is no widely usable way to create mission-realistic attack data. Generic IT labs and pure simulation miss ground–link–space timing, radio effects, and operational modes; export controls and proprietary interfaces further restrict sharing and instrumentation. The result is a shortage of trustworthy datasets for studying adversary TTPs, validating detectors, and training operators. The remedy is a modular testbed strategy comprised of digital twins with selectively inserted hardware-in-the-loop driven by the question under test and instrumented to emit synchronized command/telemetry, process/file, memory-integrity, and bus/link traces. Synthetic data should be generated from these twins with explicit provenance so results are comparable across teams.

Communication and cryptography issues dominate the second cluster of problems. Delay-/disruption-tolerant operation breaks assumptions about freshness and ordering, making key establishment, revocation, and replay defenses fragile on legacy waveforms. Post-quantum cryptography must be planned at the protocol level, not patched in, and optical/QKD concepts need evaluation against pointing loss, weather, and scheduling realities. Internally, many space system platforms remain flat: subsystems share buses without message-level authentication or authorization. Moving toward zero-trust within the vehicle and standardizing minimal, interoperable logging and attestation would close recurring gaps. In parallel, link protection must explicitly address *jamming and spoofing* with sensing-and-mitigation loops and *adaptive* RF/optical protocols that maintain integrity under Doppler, scintillation, and variable contact geometry, and key management and trust must operate across ground–link–space with DTN-aware revocation/rekey and onboard entropy/key health checks.

Detection, autonomy, and response form the third cluster. AI/ML anomaly detection faces sparse labels and physics-induced artifacts (radiation, Doppler, eclipse power transients) that mimic attacks; onboard compute and energy limits constrain model size and update cadence; explainability is required for autonomous action. Beyond security analytics, *predictive maintenance* on flight subsystems and *AI-based threat-intelligence fusion* for space telemetry are needed to anticipate degradations and prioritize hunts. Today's resilience mechanisms often erase the very evidence needed for attribution. Systems, therefore, need provenance-preserving ECC/TMR and scrubbing, append-only anomaly journaling that survives resets, and downlink strategies that trickle forensic records over multiple passes. Response playbooks must assume intermittent contact: contain while preserving observability and commandability, re-key under DTN constraints, and execute ranked recoveries that prioritize mission-critical services.

Finally, space's cyber-physical character and governance context raise problems that tools alone cannot solve. Security assessment must fuse cyber telemetry with SSA to reason about proximity operations, illumination changes, and constellation effects; redundancy and graceful degradation must preserve control and downlink rather than merely "stay on." For long missions, *on-orbit servicing and manufacturing (OSAM)* should be planned as controlled security touchpoints, and *secure IoT/edge nodes* treated as first-class participants in command and sensing. Constellation behaviors also introduce *swarming attack/defense* dynamics, requiring coordinated detection and topology-aware degradation. Supply-chain assurance, standardized secure-by-design stacks (measured boot, crypto agility), and explicit end-of-life/serviceability paths are prerequisites for long missions. Policy and standards remain fragmented – liability across commercial/government assets is unclear, and sharing is constrained, so progress depends on harmonized norms for TT&C protection, minimal common data schemas, and adoption of emerging technologies (confidential computing, PQC, quantum/optical links) only when backed by mission-level threat models and viable update paths. Adoption decisions should further consider *dynamic payload adaptability* and *AI-assisted space-traffic management*, each gated by partitioning, attestation, and robust update mechanisms. Environmental extremes, orbital dynamics, and irretrievability make early design choices tricky; getting these foundations right is the only scalable risk reducer. The details of identified issues are listed in Table 1.

Table 1 Consolidated challenges identified by participants.

Category	**Identified Problems and Topics**
Cyber Range and Simulation	Realistic cyber range simulations; High-fidelity test environments; Digital twins with selective hardware-in-the-loop; Scenario-based threat/attack simulation; Virtualization and emulation; Modular/adaptable testbeds; Operator training and awareness scenarios; Attack modeling and adversary emulation; Synthetic data generation with provenance
Delay/Disruption-Tolerant Networking (DTN)	Store–carry–forward security; Freshness/ordering under long delays; Contact scheduling effects; Robust replay protection and expiry; Routing and identity under fragmentation; DTN-aware revocation/rekey; Protocol interoperability across DSN/cislunar contexts
Secure Communications and Encryption	PQC (algorithms and *protocols*); QKD/optical feasibility and operations; Robust RF/optical authentication; Jamming/spoofing detection and response; Crypto for ground-space links under high BER/Doppler; Link-layer vs. end-to-end protections
Key Management & Trust Infrastructure	Mission-phase keying (commissioning, cruise, critical ops); Key distribution across ground–link–space; Compromise recovery and re-bootstrap under DTN; HSM/TEE use on ground and onboard; Entropy health and key/credential aging in radiation environments
AI and Autonomous Cybersecurity	AI/ML intrusion detection; Physics-conditioned anomaly detection; Onboard constraints (compute/energy/update cadence); Explainable autonomy and fail-safe action; Predictive maintenance vs. adversarial ML risks; AI-based threat intelligence
Secure-by-Design and Hardware Security	Standardized but diversified stacks; Internal message-level authz/authn (zero-trust within spacecraft); Measured/secure boot; Firmware integrity and updateability (A/B, shadow execute); Embedded crypto modules; Supply-chain security and component provenance
Incident Response and Forensics	DTN-compatible incident playbooks; Autonomous containment that preserves commandability/observability; Provenance-preserving ECC/TMR/scrubbing; Append-only anomaly journals surviving resets; Preplanned recovery options and evaluation
Cyber-Physical Resilience & SSA Coupling	Cyber with SSA (proximity operations, illumination changes); Constellation-level behaviors; Hybrid physical–cyber assessment; Redundancy and graceful degradation preserving downlink/control; Secure IoT/edge nodes in space
Safety–Security Co-engineering & Assurance	Composability of controls across EPS/ADCS/TT&C/payload; V&V for mixed safety–security requirements; Certification and testing under space constraints; Formal interface contracts to avoid harmful emergent behavior
System-of-Systems & Federated Operations	Multi-operator constellations; Cross-domain data sharing; Inter-organisational trust, SLAs, and liability; Mission handover and coalition operations
Policy, Governance, and Standards	International standards and interoperability (CCSDS/DTN/PQC-ready); Threat-intel sharing; Export controls and proprietary interfaces limiting instrumentation and reproducibility; Liability across commercial/government assets; Secure software/hardware supply-chain practices
Emerging Technologies and Trends	Digital twins for vulnerability testing (declared fidelity/limits); Dynamic payloads and modular experimentation; AI-assisted traffic management; Quantum communications; Confidential computing/TEEs
Aerospace Programmatics & Infrastructure	Launcher/ground infrastructure dependencies; Rideshare/hosted-payload risks; Power/propulsion/peripherals constraints; Mission cost/entry barriers
Unique Space Environment Challenges	Radiation, thermal cycling, micrometeoroids; Orbital dynamics and contact geometry; Irretrievability and limited servicing; Communications delay/intermittency; Environment-induced ambiguity that complicates attribution

5 Working Groups

5.1 Seminar Organization

In the afternoon of the first day of the seminar, the participants decided to divide the working groups into four teams: Prepare/Attack, Protect, Detect, and Respond. Each group started by identifying the top three pressing issues within its respective group based on the identified open problems in Table 1.

5.2 Working Group on Attack/Prepare

The Attack/Prepare group opened by enumerating blockers to credible attack research against space systems. Three roots emerged. First, an evidence deficit: there are no trustworthy, shareable attack datasets aligned with mission context. Second, legal and contractual barriers: export controls, proprietary interfaces, and vendor NDAs limit sharing, instrumentation, and reproducibility. Third, a fidelity gap: the coupling of ground–link–space timing, radios, and mission logic means generic IT labs and pure simulation fail to capture the observables that matter for adversary study.

On that basis, the group specified what a study environment must produce: (i) precondition metadata (architecture, communications characteristics, software/firmware, and access-control surfaces), (ii) nominal mission traces for the same surfaces, and (iii) aligned attack traces. Existing technique catalogues do not provide worked implementations with synchronized metadata and traces; only an integrated environment can generate all three coherently across the chain.

The resulting outcome was a testbed/digital-twin workflow rather than a stand-alone simulator. Fidelity is chosen by the question under test; hardware-in-the-loop is used where it changes observables (e.g., C&DH, EPS, ADCS, radio/SDR paths); environmental context (eclipses, radiation belts, Doppler) is modeled to shape timing and errors; and minimal instrumentation is standardized so different teams can build threat-driven twins yet still yield comparable datasets. Short-term actions recorded by the group include surveying existing flatsats and ranges, defining the instrumentation and data schemas up front, and packaging adversary scenarios mapped to space-relevant TTP catalogues for reproducible execution.

5.3 Working Group on Detect

The Detect group treated spacecraft and ground detection as a coupled problem under sparse observability and DTN. It catalogued the data required for practical methods, provenance-rich command/telemetry (counters, origin, timing, mode), process and file events, memory-integrity evidence, and internal bus/link signals, and drafted machine-actionable examples to enable sharing (e.g., command-origin deviations during specific modes; star-tracker reference-hash mismatches; mode–file/process inconsistencies). The group documented why conventional IDS tooling underperforms on mission traffic: freshness and ordering are probabilistic, error bursts and Doppler shifts mimic adversarial behaviour, and semantics are mission-specific.

Outcomes included evaluation expectations and forensic-readiness requirements. Detectors should condition scoring on physics (SAA passages, eclipses, space-weather episodes), separate environmental from adversarial false alarms, and be compared on corpora created in the

Attack/Prepare testbed. Resilience must not erase evidence: corrections and scrubbing events are to be provenance-preserving; anomaly journals must survive safe-mode resets; and downlink strategies must support trickle transmission over multiple passes.

Finally, the group recorded a range design specifically for detection research: start from clear objectives (onboard vs. ground focus), derive fidelity from those objectives, instrument at the points that expose adversary behaviour, and include 0-constellation and deep-space cases so identity, routing, and delay artefacts are exercised in a controlled way.

5.4 Working Group on Protect

The Protect group concentrated on architectural measures that hold over long missions and constrained update paths. Recurrent sources of risk were distilled from rideshare/hosted-payload arrangements, evolving network topologies (DSN/DTN, cislunar), standards and legacy components, and "X-as-a-service" ground operations. The baseline recorded by the group comprises strict internal segmentation with message-level authentication/authorization, measured boot with dependable key management and crypto agility (including PQC transition planning and re-key under DTN), and *updateability as a security requirement* (A/B images, shadow execution with telemetry-backed equivalence before commit).

Legacy integration was treated directly: risk cannot be eliminated by isolation alone. The working group specified service wrappers that enforce modern controls around older radios/payloads, dependency-longevity planning and spares, and explicit end-of-life options in contracts. A closing thread addressed composability: safety and security controls must be engineered so interactions across EPS, ADCS, TT&C, and payloads are predictable, with configuration governance to prevent harmful emergent behaviour. The seminar distinction was kept explicit: resilience restores function; protection must also preserve the truth about causes.

5.5 Working Group on Respond

The Respond group produced an operational playbook aligned with mission assurance. Preparation and monitoring come first (simulations, validated backups, rehearsed safing procedures). Identification focuses on locating adversary activity across the ground–link–space chain, understanding mechanisms and privileges (including misuse of legitimate tooling), and protecting time-critical services. Isolation is defined as containment that keeps observability and commandability intact.

Immediate recovery proceeds by ranked options, revocation and re-keying, surgical subsystem shutdowns, and only then broader isolation while assessment continues. Longer-horizon recovery restores reachable systems and stands up replacements when assets are unreachable. The cycle closes with learning and evaluation: timelines, costs, and decisions are documented, and specific hardening actions feed back into Protect and Detect. Throughout, actions are chosen to remain safe if symptoms are environmental and to reduce manipulability if they are adversarial, and all steps maintain an audit trail robust to resets and fragmented downlinks.

6 What Makes Space So Different

Addressing space cybersecurity requires a paradigm shift as the challenges are not incremental extensions of terrestrial problems, but represent a qualitative leap in complexity and nature, arising directly from the unique environment, constraints, and operational dynamics of space. At Dagstuhl, we consolidated these challenges into three foundational categories that define why cybersecurity for space assets must be treated differently:

Challenge 1: Space Environment Physical Constraints

Spacecraft operate in an environment defined by physical extremes, unlike any terrestrial system. Radiation is particularly critical: spacecraft are exposed to a complex mix of high-energy charged particles, protons, heavy ions, and electrons from solar, galactic, and extragalactic sources. While missions in Low-Earth Orbit (LEO) benefit from partial shielding, those in higher or interplanetary orbits face far more severe and sustained radiation conditions.

Radiation induces both transient and permanent effects on electronics, including bit flips, logic faults, and cumulative degradation. While these are long-recognized reliability issues, their unpredictable nature also complicates cybersecurity. A single anomaly may be environmental or adversarial, and traditional fault-tolerance techniques such as Triple Modular Redundancy (TMR) or memory checksums restore functionality without questioning causality. As a result, spacecraft may recover from a disruption yet remain blind to whether it originated from natural radiation or deliberate manipulation. This strategic ambiguity gives adversaries plausible cover: disruptions coinciding with solar flares or radiation belt passages may be dismissed as environmental, allowing targeted attacks to masquerade as background noise.

The same uncertainty extends beyond onboard systems to spacecraft communications. Space-to-ground and inter-satellite links face high latency, limited bandwidth, and intermittent availability. Corrupted packets, dropped sessions, or protocol desynchronization may result from Doppler shifts or radiation, but also from replay, delay, or spoofing attacks. In deep space, where space weather forecasting is uncertain and real-time environmental telemetry is limited, distinguishing the two is especially difficult.

Even cryptographic mechanisms are vulnerable to this ambiguity: radiation-induced bit flips in keys or entropy pools may manifest as failed authentication, broken sessions, or malformed messages, while symptoms are indistinguishable from malicious tampering. Thus, both system and communication layers face the same foundational challenge: defending against adversaries in an environment where natural effects can always provide plausible deniability.

While radiation provides the most direct cybersecurity concern, other environmental extremes reinforce the same ambiguity. Thermal cycling can shift timing margins, accelerate component aging, and degrade entropy sources or key storage, producing effects that resemble active tampering. Vacuum-driven outgassing and material fatigue, as well as sporadic micro-meteoroid or debris impacts, primarily threaten reliability but can manifest as resets, sensor drift, or link interruptions that mimic denial-of-service or integrity attacks. In contested settings, such anomalies offer adversaries plausible cover: without physics-informed diagnostics, operators may misclassify malicious interference as natural degradation.

Challenge 2: System Isolation

Space missions operate under a condition of absolute isolation: after launch, hardware can never be retrieved or replaced, and the mission must unfold with the systems committed at liftoff. Spacecraft cannot undergo hardware servicing, trusted forensic inspection, or manual reset. Their only external visibility comes through narrow telemetry channels, and any repair or mitigation must rely on pre-installed onboard logic or constrained, high-risk command uplinks from the ground. The permanence of these constraints is magnified by mission lifespans: probes such as Voyager have remained operational for nearly half a century without physical maintenance, underscoring how design choices made before launch must endure for the full mission lifetime.

A useful comparison is with industrial control systems such as chemical plants. In these settings, the process is the mission, but the cyber-physical control layer remains serviceable: controllers can be replaced, sensors recalibrated, and unit operations re-engineered during maintenance windows while the underlying process continues. By contrast, a spacecraft fuses mission and controller into a single, unreachable asset: its trajectory, sensing geometry, power and thermal envelope, and actuation topology together constitute the mission and cannot be separated from it once deployed.

These properties have concrete security implications. Assurance becomes effectively one-shot: vulnerabilities or misconfigurations that escape pre-launch detection may remain exploitable for years or decades. Monitoring and incident response are limited to the mechanisms designed from the outset. Recovery depends on autonomous mechanisms whose correctness and robustness are themselves part of the attack surface. Certification and trustworthiness, therefore, evolve differently in orbit: security is reinforced not by just periodic patching or audits, but by sustaining resilience in isolation over the mission's full operational life.

Challenge 3: Autonomy Under Extreme Latency

Even when a spacecraft remains functional, communication is constrained by distance and orbital dynamics. Deep space missions experience round-trip latencies of tens of minutes or more, and even low Earth orbit missions can encounter extended blackouts due to orbital dynamics, power constraints, or interference. In such environments, autonomy is not optional but operationally required. Yet autonomy, when combined with extreme communication delay, introduces a distinct class of security challenges.

From a security perspective, autonomy under extreme latency means the spacecraft must serve as its own guardian, at least intermittently. With no possibility for timely human verification, it must assess its state, detect attacks and anomalies, and respond to threats locally and in real-time. Traditional system monitoring mechanisms such as watchdog timers, hardware redundancy, and fail-safe control modes assume that anomalies can eventually be observed or reset through external intervention. Yet in autonomous settings, this assumption does not hold. These mechanisms respond to symptoms, not causes, and are typically agnostic to adversarial intent. A spacecraft may suffer degradation not as a result of accidental faults, but due to subtle manipulation of internal behavior. Partial corruption of command parsing, sensor fusion, or actuator logic may evade fault detection entirely while causing long-term damage. This is particularly dangerous when the degradation affects system-wide processes such as thermal regulation, power control, or attitude adjustment.

For example, a denial-of-service condition exploiting algorithmic complexity, such as a ReDoS (Regular Expression Denial of Service) attack, could induce excessive CPU or bus contention during thermally critical mission phases. If this delays or suppresses heater activation, the spacecraft may cool below its operational threshold, preventing battery bootstrap and potentially pushing components outside of their specified tolerances. Under autonomous operation, such faults may not be correctly attributed or mitigated in time, leading to cumulative and unrecoverable subsystem degradation or a safe mode configuration that is less resilient than the nominal configuration, opening up additional attack vectors.

Historical incidents show how physical degradation, even when unintentional, can result in irreversible failure. The ROSAT satellite [4], for instance, was equipped with a highly sensitive X-ray telescope that required its detectors to remain covered when not in use. However, due to a software misconfiguration in its attitude control system, the telescope was inadvertently pointed directly at the Sun during an operational maneuver. The onboard logic failed to issue a shutdown, leading to the destruction of the sensor due to solar overexposure [4]. This incident highlights how inadequate safeguards, under autonomous conditions, can lead to catastrophic outcomes from entirely foreseeable edge cases.

In contrast, the Stuxnet malware demonstrated that adversaries can deliberately induce long-term mechanical damage while concealing intent [6]. Stuxnet targeted industrial control systems running on Siemens S7-300 PLCs used in Iran's Natanz uranium enrichment facility. The attack specifically manipulated the rotational frequency of gas centrifuges used to separate uranium isotopes. These centrifuges were designed to operate within a narrow frequency window, typically around 1,064 Hz. Stuxnet intermittently altered this frequency, forcing the centrifuges to accelerate far beyond their nominal speed (reportedly up to 1,410 Hz) and then rapidly decelerate or oscillate unpredictably. These deviations were brief enough to avoid immediate failure but frequent and severe enough to create cumulative mechanical fatigue, misalign rotor assemblies, and eventually cause bearing damage or rupture.

Looking at these two cases, we argue that autonomous security must therefore operate under conditions of incomplete information, degraded sensing, and evolving mission context. The system must reason not only about whether a behavior is faulty, but also whether it is plausible given its location, trajectory, power state, and subsystem interaction, and provide its reasoning to operators when there is a connection window available for further verification.

Additionally, extreme latency and autonomy disrupt core assumptions about identity, authentication, and message integrity. Protocols designed for synchronous or near-real-time networks, such as challenge-response, key renegotiation, or session handshakes, become infeasible. Communication delays, link outages, and high bit-error rates mean that space networks must adopt Delay-tolerant Networking (DTN) principles, where messages are asynchronously relayed, buffered, and reassembled. However, DTN conditions undermine traditional guarantees of freshness, liveness, and ordering primitives on which most terrestrial cryptographic protocols depend.

The result is a security model in which authenticity and trust are probabilistic rather than deterministic. For example, a packet received during an expected contact window, from a plausible antenna orientation, and with correct signal power may be considered more trustworthy than one that deviates from these constraints. Yet this judgment must be made onboard, in real time, with limited context, without external validation, and with all existing power and processing budget limitations.

Challenge 4: Pervasive Exposure and Asymmetric Attack Surface

Space systems possess a fundamentally different attack surface from terrestrial systems, not just in extent but in asymmetry, persistence, and observability. The attack surface spans physical, cyber, and RF domains, each with unique entry points and defense limitations. What sets this domain apart is not the existence of more vectors, but the inability to constrain, observe, or attribute many classes of attacks.

From a security perspective, spacecraft are highly integrated cyber-physical platforms with interconnected subsystems. Each of these may become an attack vector or fault amplifier. For example, access to a thermal management controller or a fault handler may provide indirect control over avionics or memory protection logic. Many existing spacecraft architectures use shared communication buses and unsegmented internal channels, meaning that subsystems can exchange messages over a common interface without isolation or message-level authentication. These designs were historically justified by the assumption that spacecraft are physically inaccessible to adversaries, and therefore internal communications would remain trustworthy and uncontested. This assumption no longer holds in a world where remote code execution, protocol exploitation, or malicious payload injection can be initiated from Earth.

Moreover, modern spacecraft increasingly incorporate commercial off-the-shelf components, third-party software, and open-source libraries [7]. These introduce opaque and often unverified dependencies into mission-critical systems. Vulnerabilities in telemetry handlers, decompression modules, or firmware may go unnoticed until operational deployment, and most spacecraft cannot fully patch or revalidate such components post-launch [8, 3].

Additionally, the most persistent and unavoidable exposure lies in the continuous use of radio frequency or optical interfaces for communication. Spacecraft must maintain always-on RF or optical interfaces for telecommands and data operations. These channels are predictable in time and frequency and inherently exposed.

Additionally, the ground segment introduces a systemic and often underestimated vulnerability. Ground control software, mission scheduling systems, and telemetry storage infrastructure can be compromised to influence spacecraft indirectly. The 2022 Viasat KA-SAT attack is a relevant example [1], where satellite communications were disrupted at scale without modifying satellite firmware, illustrating that terrestrial infrastructure remains a viable entry point.

The asymmetry is further exacerbated by the imbalance between attackers and defenders. Attackers can observe orbital paths, predict visibility windows, and time attacks precisely. Defenders, in contrast, often operate with outdated or intermittent telemetry, lack real-time access, and have little visibility into the presence or behavior of an attacker.

A new facet of this asymmetry is the use of spacecraft to target or probe other spacecraft directly. Nation-state actors have increasingly conducted proximity operations, where one satellite shadows or approaches another to observe its behavior, gather RF emissions or assess response patterns. These interactions, often described as on-orbit reconnaissance or Rendezvous and Proximity Operations (RPO), blur the line between surveillance and preparatory attack, especially when used to map out operational vulnerabilities or test defense thresholds. For example, in 2020, a Russian satellite (Kosmos 2542) maneuvered close to a U.S. military satellite (USA 245), prompting public warnings from U.S. Space Command about potential antisatellite behavior [5]. These actions are rarely transparent, difficult to verify, and often designed to remain below escalation thresholds. As space becomes more contested, inter-spacecraft operations may evolve into a common vector, requiring new security models that consider not only terrestrial threats but also orbital adversaries.

6.1 Summary

Taken together, these challenges show that space cybersecurity is fundamentally unlike any other cyber-physical security problem. Other domains may share fragments of these issues, but only in space do they converge inseparably and persist for the entire mission lifetime. Harsh physical environments, absolute isolation, extreme communication delays, and pervasive exposure do not just complicate defense; they redefine it. The result is an attack surface that is broader, more persistent, and less observable than in any terrestrial system, forcing defenders to treat incomplete information, degraded sensing, intermittent trust, and adversarial uncertainty as normal operating conditions.

This conclusion reflects the consensus of the Dagstuhl Seminar, where more than forty experts from the space and cybersecurity domains, including specialists in cyber-physical systems, concluded that space constitutes a qualitatively different security environment. Meeting these challenges requires more than adapting terrestrial techniques: it demands fundamentally new approaches that embed physics-informed reasoning, resilience without servicing, and autonomy designed to operate securely under uncertainty for decades at a time.

7 Future Work and Challenges Ahead

The rapid evolution of the space domain, characterized by the proliferation of commercial mega-constellations, increasing autonomy, the steep increase in data throughput capacity, and the extension of terrestrial networking paradigms into orbit, presents a complex and dynamic cybersecurity landscape. The discussions at the Dagstuhl Seminar crystallized a consensus that future research and development must move beyond traditional security research and consider the specificities of the space environment and its constraints. This section outlines the critical frontiers and formidable challenges identified by the seminar participants during the fourth and fifth days of the seminar, providing a roadmap for the academic, industrial, and governmental efforts required to secure the future of space operations. The challenges involve and merge many disciplines in security research, from foundational cryptographic transitions to the complexities of autonomous defense, secure hardware design, cyber-physical resilience, and the establishment of robust international governance.

7.1 Secure Space Communications and Encryption in the Quantum Era

The looming threat of fault-tolerant quantum computers capable of breaking current public-key cryptography (e.g., RSA, ECC) using algorithms like Shor's necessitates a fundamental overhaul of cryptographic systems for space. This is not a distant, theoretical concern but an urgent operational reality. Given the long lifecycles of spacecraft, which can operate for 15–20 years, and the general impossibility of post-launch hardware upgrades, a proactive transition to quantum-resistant security is not merely advisable but mission-critical. Systems launched today with vulnerable cryptography could have their communications intercepted and stored, ready to be decrypted by a future quantum computer, a "harvest now, decrypt later" attack that poses an unacceptable risk to long-term national security and commercial intellectual property. This challenge bifurcates into two primary, and often complementary, research avenues: the near-term deployment of Post-Quantum Cryptography (PQC) and the long-term, ambitious development of Quantum Key Distribution (QKD) for space applications.

7.1.1 The Duality of Quantum Key Distribution (QKD) and Post-Quantum Cryptography (PQC)

The path toward quantum-resistant space systems is defined by the distinct characteristics, trade-offs, and timelines of QKD and PQC. Understanding this duality is fundamental to developing a coherent security strategy.

7.1.1.1 The Promise and Peril of QKD

Quantum Key Distribution (QKD) represents a paradigm shift in secure communications. Its security guarantee is not based on the presumed computational difficulty of a mathematical problem, but on the fundamental laws of quantum physics, such as the no-cloning theorem and Heisenberg's uncertainty principle. This provides information-theoretic security, meaning that an eavesdropper's attempt to intercept and measure the quantum states (e.g., polarized photons) used to generate a key would inevitably disturb the system, revealing their presence to the legitimate parties. In theory, this makes the key exchange impervious to any future advances in computing, including quantum computers.

The feasibility of this technology for space has been convincingly demonstrated. Experiments, most notably China's Micius satellite launched in 2016, have successfully established space-to-ground and inter-satellite QKD links, distributing secure keys over distances exceeding 1,200 km. These missions proved that satellite-based QKD can overcome the distance limitations of terrestrial fiber-optic QKD, which suffers from exponential signal loss, and could form the basis of a future global quantum internet.

However, despite its theoretical promise, QKD faces immense practical challenges that currently limit its widespread deployment. Firstly, QKD is only a partial solution; it secures the distribution of a symmetric key but does not provide authentication. The source of the QKD transmission must be authenticated using classical methods, which today means relying on pre-placed keys or, ironically, PQC, making QKD vulnerable to man-in-the-middle attacks if the authentication layer is weak. Secondly, QKD requires specialized, costly, and inflexible hardware, such as single-photon detectors and precise pointing systems, which are difficult to integrate into satellite buses and impossible to upgrade post-launch. Thirdly, free-space optical links are susceptible to disruption from atmospheric conditions like turbulence and cloud cover, and current key generation rates remain far too low for high-bandwidth applications, with some demonstrations yielding only a few bits of secure key per satellite pass. Finally, the theoretical security of the protocol can be undermined by practical side-channel attacks that exploit imperfections in the physical hardware, and its inherent sensitivity to any disturbance makes it highly susceptible to denial-of-service (DoS) attacks. These significant limitations have led to a cautious stance from security bodies like the U.S. National Security Agency (NSA), which currently does not recommend QKD for securing National Security Systems until these fundamental implementation and security validation challenges are overcome.

7.1.1.2 The Pragmatism of PQC

Post-Quantum Cryptography (PQC) offers a more immediate and pragmatic path to quantum resistance. PQC algorithms are classical, meaning they can run on existing computer hardware, but are based on mathematical problems, such as those found in lattice-based, hash-based, or code-based cryptography, that are believed to be computationally infeasible to solve for both classical and quantum computers.

A major driver for PQC adoption is the progress in standardization. The U.S. National Institute of Standards and Technology (NIST) has completed its multi-year competition and has finalized the first standards for PQC algorithms. These include CRYSTALS-Kyber (standardized as ML-KEM) for key encapsulation and CRYSTALS-Dilithium (standardized as ML-DSA) for digital signatures, providing a vetted foundation for industry to build upon. This has spurred active development, with projects already underway by organizations like the European Space Agency (ESA) to design and implement PQC-based cryptographic systems for securing satellite telecommunication applications, particularly command and control links. Furthermore, space standardisation organisations such as the Consultative Committee for Space Data Systems (CCSDS), are adopting the NIST recommendations already.

PQC is not, however, a simple drop-in replacement for current cryptographic standards. Its security remains computational, not absolute, and the field is still maturing. Several PQC candidate algorithms, including some that reached advanced stages of the NIST process, have been broken by subsequent cryptanalysis using classical computers, highlighting the potential for future vulnerabilities to be discovered. For space systems, the most pressing challenges are practical. PQC algorithms often require significantly larger key sizes and signatures, and are more computationally intensive than their classical counterparts. This poses a substantial problem for the highly constrained Size, Weight, and Power (SWaP) environment of satellites, where processing power and bandwidth are scarce resources. Furthermore, the harsh radiation environment of space introduces the risk of Single Event Upsets (SEUs), bit-flips caused by cosmic rays, which could corrupt complex PQC calculations, potentially leading to authentication failures or security breaches. This makes the research and development of fault-tolerant PQC implementations (which would not necessarily need to depend on expensive radiation-hardened components), likely involving specialized hardware and error-correcting codes, a critical area for future work.

7.1.1.3 A Hybrid and Risk-Stratified Future

The ongoing debate is not a simple choice of "QKD vs. PQC." Rather, the evidence points toward a future where the two technologies are integrated into a hybrid, risk-stratified architecture. PQC is the only viable path for achieving broad crypto-agility in the near term. Its software-based nature allows it to be deployed on existing and new systems to secure the vast majority of commercial and tactical communications. It will become the workhorse of space cryptography.

However, PQC alone cannot defend against the "harvest now, decrypt later" threat for data that requires confidentiality for decades or longer. This is where QKD finds its crucial niche. A hybrid model is therefore necessary. In this model, PQC provides the robust, authenticated channel required for QKD to operate securely, protecting it from man-in-the-middle attacks. QKD, in turn, provides an information-theoretically secure method for distributing keys for the highest-value, strategic communication links where the cost and complexity are justified. This could include, for example, securing command links for national security satellites, establishing a secure key-exchange backbone for deep space missions, or protecting critical diplomatic communications.

This leads to a tiered cybersecurity model for space assets. High-value, state-owned strategic assets may be equipped with expensive, hardware-based QKD systems for ultimate long-term security. In contrast, large commercial constellations, where cost is a primary driver, will rely on more agile but computationally-based PQC. This inevitable stratification will create profound new challenges for interoperability between different security domains, for the development of international policy, and for the creation of standards, as a single definition of "secure" will no longer apply universally across the space ecosystem.

Table 2 contrasts Quantum Key Distribution (QKD) with Post-Quantum Cryptography (PQC) for satellite links. QKD provides information-theoretic security but demands specialized optics and precise pointing, yields limited/fragile key rates, and supplies key distribution only (no native authentication). PQC relies on computational hardness yet is software-deployable on existing hardware, already standardized (e.g., ML-KEM/ML-DSA), and supports both key establishment and digital signatures. In practice, PQC is the default for securing command and control, while QKD is complementary for bulk key pre-distribution where optical links and SWaP budgets permit.

Table 2 Comparative Analysis of QKD and PQC for Satellite Communications.

Attribute	**Quantum Key Distribution (QKD)**	**Post-Quantum Cryptography (PQC)**
Security Basis	Information-theoretic, based on laws of physics; provides forward secrecy against future computational advances.	Computational, based on hardness assumptions believed to resist quantum attacks.
Maturity (TRL)	Low to medium; experimental demonstrations (e.g., Micius) are successful but not yet mature for broad space deployment.	Medium to high; NIST standards exist (e.g., ML-KEM, ML-DSA); space-grade implementations are in development.
Implementation	Hardware-intensive; requires single-photon sources/detectors and precision pointing; not software deployable.	Software-based; runs on existing hardware, deployable via firmware or software updates.
Primary Function	Key distribution only; authentication must be provided separately.	Key encapsulation and digital signatures for authentication.
Suitability for C&C	Challenging; low key rates and DoS susceptibility limit use for critical command links; useful for bulk key pre-distribution.	High; suitable for securing command and control links due to software nature and authentication support.
Key Vulnerabilities	Side-channel attacks on optics/electronics; DoS from atmospheric or malicious interference; lack of built-in authentication.	Potential future cryptanalytic breaks; implementation bugs; software side channels.
SWaP Impact	High; adds dedicated hardware with mass, power, and volume penalties.	Moderate; higher compute and memory than classical crypto, but no new hardware subsystem required.
Fault Tolerance	Highly sensitive to disturbances, atmospheric effects, and pointing errors.	Complex algorithms susceptible to SEUs in radiation, requiring fault-tolerant design.

7.1.2 Securing High-Bandwidth Optical and RF Communications

The transition to high-throughput communication links, particularly optical/laser inter-satellite links (ISLs) and space-to-ground connections, is a key enabler for future space services, from global broadband to massive deep-space science and Earth observation downlinks. While offering unprecedented bandwidth, these links also present high-value targets for sophisticated adversaries seeking to conduct eavesdropping, jamming, or spoofing attacks. Securing these communications on all layers (physical, link, and network) is a critical challenge.

Future research must focus on developing advanced defense mechanisms tailored to the unique physics of these channels. For optical communications, this means moving beyond defenses designed for RF systems. Research is needed into techniques that can distinguish malicious jamming or spoofing from natural environmental interference, such as atmospheric scintillation, which can cause similar signal degradation. AI-based signal analysis, capable of learning the subtle signatures of both atmospheric conditions and deliberate attacks, presents a promising avenue for robust detection.

Furthermore, as satellites become nodes in large, interconnected mesh networks, resilience cannot depend on a single point-to-point link. Future work must involve the design of secure and resilient routing protocols for these large-scale optical constellations. Such protocols must be able to detect a compromised or failed node and dynamically re-route traffic through trusted paths, ensuring network availability and graceful degradation of service rather than catastrophic failure. This also requires the development of novel signal transformation and Transmission Security (TRANSEC) protocols that enhance resilience against interception and manipulation at the lowest layers of the communication stack.

7.1.3 Authenticated and Resilient Command & Control (C&C)

The command and control (C&C) link is the umbilical cord to a spacecraft; its compromise can lead to the partial or total loss of the asset. Securing this link requires a multi-layered, end-to-end approach that extends from the operator at a control center to the satellite bus in orbit.

A key area for future work is the development of next-generation onboard defenses. This involves moving beyond static defenses to adaptive, space-specific firewalls and on-bard Intrusion Detection Systems (IDS). These systems must be capable of operating effectively within the severe resource constraints of a satellite's onboard computer, analyzing command flows and telemetry for signs of unauthorized activity.

Equally important is ensuring the integrity of the entire C&C chain. An attack is just as likely to originate from a compromised ground segment as it is to target the space-to-ground link. Therefore, cryptographic security, likely based on the new PQC standards, must be implemented and rigorously verified across all components of the system, including ground station software, network infrastructure, and operator terminals. This ensures end-to-end trust and prevents an attacker from bypassing space link encryption by compromising a vulnerable terrestrial component. It also needs to include formal security proofs for space-link communication security standards such as the CCSDS Space Data-Link Layer Security (SDLS) standard familiy.

As the ground segment state-of-the-art moves more into the shared infrastructure approach (e.g. ground station as a service, ground segment as a service) and multi-mission support (one ground segment for many missions) ground segment resilience remains a key aspect with elements such as multi-mission zero-trust architectures taking a major role in research.

7.2 AI-Driven and Autonomous Space Cybersecurity

As space missions venture further from Earth into deep space and constellations grow to encompass thousands of interconnected nodes, the operational paradigm is fundamentally changing. The signal propagation delays, which can range from minutes to hours for deep space missions, combined with the sheer scale and complexity of mega-constellations, make direct human-in-the-loop control for cybersecurity functions untenable. In this new era, autonomy is not an option but a necessity. The central challenge for the research community is to develop Artificial Intelligence (AI)-driven systems that can autonomously detect, reason about, and respond to cyber threats in real-time. The ultimate goal is to create self-defending space assets capable of ensuring their own survival and mission success without constant human intervention.

7.2.1 AI for Advanced Intrusion Detection and Response

Traditional security tools, such as signature-based Intrusion Detection Systems (IDS), are fundamentally reactive. They are effective at identifying known threats but are easily bypassed by novel, zero-day attacks. The future of on-orbit threat detection lies in AI's ability to move beyond pattern matching to behavioral analysis. AI and Machine Learning (ML) models can be trained to establish a high-fidelity baseline of a satellite's normal operations, encompassing everything from bus telemetry and power consumption patterns to network traffic and payload activity, and then identify anomalous deviations that could indicate a compromise.

Key research directions in this area include the implementation and refinement of various AI/ML models tailored for the space environment. Unsupervised learning models like Isolation Forests and Deep Autoencoders are well-suited for anomaly detection in network traffic, while sequence-aware models like Long Short-Term Memory (LSTM) networks can detect deviations in time-series data, such as user activity logs or command sequences. Another critical application is using AI for real-time analysis of the RF spectrum. By learning the characteristics of legitimate signals, an AI system can detect and classify sophisticated jamming and spoofing attacks, providing a layer of cyber-physical defense that bridges the digital and physical domains.

However, deploying these AI models effectively presents significant challenges. A primary hurdle is managing the high rate of false positives, which can overwhelm operators and lead to alert fatigue. This is exacerbated by the inherent data imbalance in cybersecurity, where malicious events are rare compared to normal operations. Furthermore, the computational cost of complex deep learning models can be prohibitive for SWaP-constrained satellites. Finally, the “black box” nature of many AI models poses a challenge for trust and verification; operators need explainable AI (XAI) techniques to understand why an alert was triggered before they can confidently act on it.

7.2.2 Self-Defending and Self-Healing Spacecraft

Detecting a threat is only the first step; a system must also be able to respond effectively to mitigate the threat and restore its core functions. This concept of cyber resilience, the ability to withstand, operate through, and recover from an attack, is the foundation of autonomous cybersecurity.

Future work must focus on developing Tactical Autonomous Systems (TASS), which are AI-driven agents capable of executing pre-defined defensive “playbooks” in response to a detected intrusion. Upon identifying a threat, a TASS could autonomously take action, such as isolating a compromised subsystem from the main bus, rerouting critical data through trusted communication paths, or disabling non-essential functions to preserve the primary mission. In the context of large constellations, this extends to cooperative defense, where satellites can share threat intelligence and defensive postures with trusted peers. This allows the entire constellation to “learn” from an attack on a single node and collectively adapt its defenses, creating a resilient, herd-like immunity.

The ultimate goal is the creation of self-healing architectures. This involves researching systems where a satellite can not only detect and isolate a threat but also autonomously purge malware, restore critical software from a secure, read-only backup, and even apply security patches to remediate the underlying vulnerability, all without human intervention. This capability is especially critical for long-duration missions into deep space, where the extreme communication delays make interactive recovery impossible.

7.2.3 AI for Predictive Maintenance and Fault Tolerance

The line between a system fault and a cyberattack is often blurry. A physical component failure could be a precursor to a cyberattack, a vulnerability an attacker might exploit, or even the direct result of a malicious command. AI can play a crucial role in bridging the gap between traditional Fault Detection, Isolation, and Recovery (FDIR) and cybersecurity.

By applying ML models to analyze historical and real-time telemetry, AI systems can perform predictive maintenance, forecasting component failures before they occur. An unexpected prediction of failure in a healthy component could serve as an early indicator of a subtle, ongoing cyberattack. This fusion of FDIR and cybersecurity enhances overall mission assurance.

A key enabling technology for this is the concept of a digital twin. By creating a high-fidelity, physics-based virtual replica of a satellite and its environment, operators can run simulations that are impossible to conduct on the real asset. Enhanced with Generative AI, these digital twins can be used to simulate a vast range of both physical fault and cyberattack scenarios. This provides an invaluable, safe environment for training and validating AI-based detection and response models before they are deployed on the actual spacecraft, significantly improving their reliability and effectiveness. A critical challenge within this domain is the development of a trusted, automated system for deploying firmware and software patches to an orbiting satellite. Such a system is essential for both fixing bugs and remediating vulnerabilities, but it also represents a prime target for a supply chain attack, where an adversary could inject malicious code into a seemingly legitimate update. Securing this automated patching pipeline is a major research challenge.

7.2.3.1 The Double-Edged Sword of AI

While AI is a powerful enabler for autonomy, it presents a fundamental paradox: the very tools used to manage complexity introduce new, opaque attack surfaces. AI models, particularly deep neural networks, often behave as inscrutable “black boxes” whose performance is brittle under unforeseen conditions and vulnerable to adversarial manipulation, from data poisoning during training to evasion at inference. Critically, traditional software assurance methods like code review and static analysis are insufficient for these learned systems. This creates a dangerous safety-security coupling in autonomous systems like spacecraft, where a security failure (e.g., a spoofed sensor reading) can cascade into an unsafe control action with no immediate human oversight. Countering this requires a security-by-construction approach, integrating multiple layers of protection: formally specified operational envelopes to constrain actions; runtime monitors grounded in physical laws; verifiable provenance for all training data; and signed, versioned models to prevent unauthorized modification.

Table 3 summarizes how common AI/ML methods map to autonomous space-cyber tasks, from anomaly detection and self-healing control to predictive maintenance, constellation-level defense, and RF signal analysis. The techniques promise mission awareness and faster response, but face practical hurdles: scarce high-fidelity data, onboard SWaP limits, safety/verification of autonomous actions, and robustness to interference and data poisoning. Designing for explainability, forensic readiness, and secure digital-twin workflows is essential for reliable deployment.

7.2.3.2 The Rise of Agentic Adversaries

Looking ahead, attackers can field *agentic* AIs that plan, probe, and adapt with minimal human input: autonomously discovering protocol flaws, staging multi-step RF/cyber campaigns, synthesizing spoofed telemetry consistent with orbital dynamics, or timing actions to exploit DTN delays and attribution ambiguity. This lowers the barrier to persistent, tailored operations against both spacecraft and ground segments. Meeting agentic offense requires agentic defense. Beyond static detectors, we need autonomous, policy-bounded defenders that (i) reason over multi-modal evidence (telemetry, RF, ephemerides, power/thermal), (ii) enact deception, and (iii) recover safely. Practical building blocks include (but are not limited to) *agentic honeypots* (emulated subsystems, decoy services, and DTN honeynodes that absorb and fingerprint probes), moving-target defenses (keying, routing, and software diversity scheduled within power/thermal budgets), physics-aware plausibility filters (rejecting commands or state transitions that violate dynamics), and closed-loop response playbooks with human-on-the-loop authority. Integration should be split: lightweight, fail-safe agents onboard (SWaP-bounded, with hard limits and kill-switches) for fast containment; heavier agents on the ground and within digital twins for red teaming, hypothesis testing, and model retraining. All agents must ship with verifiable policies, audited action histories, and secure update/provenance channels so that an attacker cannot turn the defender into an amplifier.

Table 3 AI/ML Techniques for Autonomous Space Cybersecurity Tasks.

Cybersecurity Task	**AI/ML Technique**	**Primary Function**	**Key Challenges**
Anomaly Detection and Intrusion Response	Deep autoencoders; Isolation Forests; LSTMs	Learn a baseline of normal telemetry/network behavior and flag significant deviations as potential intrusions.	High false positives; need for high-fidelity training data; model explainability; SWaP limits for onboard inference.
Self-Healing and Autonomous Defense	Reinforcement learning; Tactical Autonomous Systems (TASS)	Isolate subsystems, reroute traffic, or disable non-essential functions to neutralize threats and restore functionality.	Reward design; safety during learning; computational cost; formal verification of autonomous actions.
Predictive Maintenance and Fault Tolerance	Supervised learning (SVM, Random Forest); Digital twins	Predict component failures from historical/real-time data and simulate faults/attacks in a virtual replica.	Distinguishing natural faults from malicious actions; twin fidelity; securing the twin itself.
Cooperative Swarm Defense	Federated learning; Swarm optimization	Train shared models across a constellation without sharing raw data and coordinate defensive maneuvers.	Communication overhead; non-IID data; robustness against poisoning from compromised nodes.
RF Signal Analysis	CNNs; Autoencoders	Detect, classify, and localize jamming/spoofing via raw spectrum pattern analysis.	Disentangling interference vs. attacks; real-time processing; large, diverse datasets.

7.3 Secure-by-Design in Space System Hardware and Software

For decades, the prevailing security model for space systems focused on protecting the communication link, treating the satellite itself as a trusted “black box” operating within a secure perimeter. This assumption is now dangerously obsolete. The advent of software-defined satellites, the increasing complexity of global supply chains, and the demonstrated potential for on-orbit malware necessitate a fundamental shift in philosophy. The principles of “secure-by-design” and “secure-by-default,” championed by bodies like the U.S. Cybersecurity and Infrastructure Security Agency (CISA), must be adopted. Security can no longer be an afterthought or a feature to be added on; it must be a foundational requirement, embedded into every layer of the system’s hardware and software from the initial design phase.

7.3.1 Hardware-Rooted Security and Trust

Software-only security measures are inherently vulnerable because they can be bypassed if the underlying hardware or boot process is compromised. To build a truly trustworthy system, trust must be anchored in immutable hardware. This applies to both the platform and payload hardware.

A critical technology for achieving this is the Trusted Platform Module (TPM). A TPM is a dedicated, tamper-resistant microcontroller that provides a hardware root of trust for critical security functions. Adapting and qualifying TPMs for the rigors of the space environment is a key area for future work. A space-grade TPM could provide a secure foundation for a satellite’s entire software stack by enabling a secure boot process, which cryptographically verifies the integrity of each piece of software before it is loaded, from the bootloader to the operating system and flight application. This ensures that only authorized, untampered code can execute. Furthermore, a TPM can provide secure key storage, protecting critical cryptographic keys from being extracted by software-based attacks, and can perform attestation, allowing the satellite to prove its identity and software state to a ground station in a cryptographically verifiable way. Research into using Physically Unclonable Functions (PUFs), such as those based on ring oscillators, can enhance attestation by creating unique, unclonable hardware “fingerprints” for each satellite.

For more computationally intensive cryptographic operations, such as those required by PQC algorithms, dedicated Hardware Security Modules (HSMs) or crypto-accelerator chips are necessary. While common in terrestrial data centers, the challenge lies in developing radiation-hardened, low-power versions of these technologies that can survive and operate reliably in the space environment.

7.3.2 Lightweight and Verifiable On-Orbit Systems

The proliferation of small satellites, particularly CubeSats, introduces a different set of challenges. These platforms have extreme SWaP constraints, which often preclude the use of traditional, resource-heavy security solutions designed for larger satellites. Securing these systems requires innovation in lightweight and efficient security.

A major research focus is the design of lightweight Intrusion Detection and Prevention Systems (IDS/IPS) tailored for resource-constrained embedded systems. This involves developing lightweight ML models, optimizing feature selection to reduce computational load, and creating distributed architectures where complex analysis and model training are offloaded to the ground segment, while a smaller, efficient inference engine runs on the satellite itself.

Another powerful approach for ensuring security in critical systems is the application of formal methods. These are mathematically rigorous techniques used to specify and verify the properties of a system. By creating a formal model of critical flight software, it is possible to mathematically prove the absence of entire classes of vulnerabilities, such as buffer overflows or race conditions, and to verify that the software behaves exactly as specified under all conditions. While historically labor-intensive, advances in tools like model checkers and SMT solvers are making formal methods more accessible. The key challenge is scaling these techniques to handle the ever-increasing complexity of modern flight software.

7.3.3 Cybersecurity for On-Orbit Servicing, Assembly, and Manufacturing (OSAM)

The emergence of OSAM and hosted payload business models fundamentally alters the security paradigm. A satellite is no longer a static, monolithic asset. Instead, the space environment is becoming a dynamic, physically interactive, and potentially multi-tenant ecosystem. This dramatically expands the attack surface and introduces novel threat vectors.

Securing robotic OSAM missions is a primary concern. An attacker who compromises the command link or autonomous logic of a robotic servicer could turn a repair mission into a deliberate kinetic attack, using the servicer to physically damage or de-orbit a target satellite. Future work must focus on securing these robotic operations, including robust authentication, encrypted command links, and verifiable autonomous logic.

These missions also create complex challenges for trust and access control. A typical servicing mission may involve multiple independent entities: the servicer owner, the client satellite owner, and potentially a third-party payload owner. This necessitates the development of new security frameworks for managing trust and access control in these multi-party interactions, including secure protocols for rendezvous, proximity operations, docking, and data exchange.

Finally, the concept of “Space-as-a-Service,” where satellite operators host third-party application code or payloads, requires robust security measures. Future research must focus on creating secure containerization or virtualization environments on satellites. These “sandboxes” must provide strong isolation to prevent a compromised payload from affecting the host satellite bus, other payloads, or accessing data it is not authorized to see.

7.3.3.1 The Supply Chain as the New Perimeter

While on-orbit security technologies like PQC and AI are critical, they can be rendered moot if a system is compromised before it ever reaches orbit. The most immediate and insidious threat vector facing the space industry today is the supply chain. A “secure-by-design” philosophy is meaningless if the components used in that design are already malicious. This elevates Cybersecurity Supply Chain Risk Management (C-SCRM) from a simple procurement issue to a primary national security challenge for space.

The logic is straightforward. Space systems are assembled from a complex, global supply chain of hardware and software components, many of which are Commercial-Off-The-Shelf (COTS) to reduce costs. An adversary can target any point in this long and often opaque chain, from injecting malicious logic into a microchip at a foundry, to inserting a backdoor into open-source software, to tampering with a component during integration. This means a satellite could be launched with a hidden vulnerability or a dormant backdoor already embedded deep within its hardware or software. Such a compromise would completely bypass all link-level encryption and on-orbit defenses, waiting to be activated by an attacker at a time of their choosing. This threat is explicitly recognized as a key concern in foundational policy documents like U.S. Space Policy Directive-5.

This reality demands a "zero-trust" approach to the supply chain itself. Future work must prioritize the development of technologies and policies to ensure the integrity of components from fabrication to launch. This includes developing and mandating cryptographically signed and verifiable hardware and software bills of materials (HBOM/SBOM), establishing programs to source critical components from trusted and vetted suppliers, and using comprehensive frameworks like the NIST Cybersecurity Framework to continuously assess and manage supply chain risk throughout a program's lifecycle, not just as a one-time check. Ultimately, assuming the supply chain can be compromised, hardware-rooted security technologies like TPMs and secure boot become the final and most critical line of defense, providing a mechanism to detect and prevent the execution of unauthorized, malicious components that may have been inserted during manufacturing. Finally, this would need to be complemented with behavior monitoring to detect changes and the use of a placed backdoor.

7.4 Cyber-Physical Resilience for Multi-Domain Missions

Cyberattacks against space systems are not merely data breaches; they are attacks on physical assets with tangible, real-world consequences. A successful cyberattack could be used to manipulate a satellite's propulsion system to alter its orbit, disable a critical Earth observation sensor during a natural disaster, or even command a satellite to perform a maneuver that causes a collision, generating a cloud of orbital debris that threatens all space activities. Therefore, future cybersecurity research must focus on the concept of cyber-physical resilience, ensuring mission continuity even when under attack, and deeply integrating cyber defense with our physical understanding of the space environment.

7.4.1 Integrating Cybersecurity with Space Situational Awareness (SSA)

Cybersecurity and Space Situational Awareness (SSA), the practice of tracking and characterizing objects in orbit, are not considered as correlated domains so far. However, a clear relationship exists and needs to be assessed further. A cyberattack can have direct physical manifestations, e.g., an unexpected maneuver or change in RF emissions, and conversely, compromised SSA data can be used as a weapon to enable a cyber or physical attack. Futhermore, cybersecurity can be used as a tool to solve the question of maneuver accountability in case of identified collision risk between two assets that are owned by different stakeholders.

Future work must focus on breaking down these silos. This requires developing a capability for "cyber-informed SSA," where cyber threat intelligence is used to guide physical monitoring. For instance, a security alert indicating a potential compromise of a satellite's command and control system should automatically trigger increased tasking of ground-based telescopes and radars to monitor that specific satellite for any anomalous physical behavior. The reverse is also true. "SSA-informed cyber defense" would use physical data as a potential indicator of a cyberattack. For example, the unexpected close approach of an unknown or non-communicative object could trigger a heightened state of cyber monitoring on the high-value asset being approached.

Furthermore, the SSA data ecosystem itself, from the global network of sensors to the data fusion centers and the operators, is a prime target for attack. If an adversary can manipulate the data that operators rely on to understand the space environment, they can cause confusion, hide their own activities, or even induce an operator to perform a disastrous "collision avoidance" maneuver against a phantom object. Research is needed to secure this entire ecosystem against data manipulation and spoofing, ensuring that operators have a trusted, verified picture of the space domain.

7.4.2 Defense Against Autonomous Swarm Threats

The miniaturization of satellites and advances in autonomous coordination are enabling the development of satellite swarms. While these swarms have many beneficial applications, they also represent a novel and potent threat. An adversarial swarm of small, inexpensive satellites could be used to conduct a distributed denial-of-service attack against a target's communication links, perform coordinated, multi-point jamming, or even physically harass or disable a high-value asset through proximity operations.

Defending against such threats requires new approaches. A key area for research is AI-driven swarm defense. This involves developing defensive AI systems that can detect, track, and predict the intent of an adversarial swarm using advanced sensor fusion and behavioral analysis. Beyond detection, future work must focus on designing friendly satellite swarms with inherent resilience. This includes architectures with decentralized command and control, bio-inspired coordination algorithms that allow for emergent, adaptive behavior, and the ability to maintain overall mission capability despite the loss of individual nodes to attack or failure. Such resilient architectures can provide a robust defense, capable of dynamically responding to and mitigating threats from adversarial swarms.

7.4.3 Cybersecurity for Deep Space Operations/ Solar-System Internet

Missions to deep space destinations, meaning Cis-Lunar and beyond, push the boundaries of operational complexity. These missions are defined by extreme communication delays induced by the light speed barrier, minutes to hours as well as frequent and predictable link disruptions due to orbital mechanics. These conditions render traditional, interactive cybersecurity protocols, which rely on real-time communication with an operations centre, inefficient and error-prone. For these missions, security must be highly autonomous and tolerant of prolonged delays and disconnection. In addition, because of the high cost of deep space missions, they are very often comprised of assets owned by different stakeholders that interoperate. This inherently raises questions of trust and routing priority and it requires fully interoperable and standardised secure communication protocols as well as a decentralized key management concept.

A foundational technology for this environment is Delay/Disruption Tolerant Networking (DTN) with the Bundle Protocol (BP) at its core. DTN is a store-and-forward network architecture designed specifically for these conditions, allowing data to be held at intermediate nodes, e.g., a Mars orbiter, until a forward link becomes available. A critical area of future work is to mature, standardize, and deploy the security protocols for DTN, such as the Bundle Protocol Security Protocol (BPSec), to provide robust confidentiality, integrity, and authentication services over these challenging links.

Beyond the network layer, onboard systems for deep space missions must be empowered to make security-critical decisions autonomously. A spacecraft cannot wait for hours for confirmation from Earth to respond to a threat. This requires the development of robust, pre-programmed security policies and advanced AI/ML capabilities that allow the spacecraft to independently assess a situation, such as whether to trust a new communication partner or how to respond to an anomalous sensor reading, and take appropriate action. This also extends to long-term key management, where new protocols are needed to securely manage cryptographic keys and handle credential revocation over mission durations that can span decades, all across high-latency communication links.

7.4.3.1 The Blurring of Lines Between Cyber and Kinetic

The emergence of physically capable autonomous systems in space, such as OSAM servicers and coordinated swarms, effectively erases the clear, traditional distinction between a “cyberattack” and a “kinetic attack.” This convergence of digital and physical threats creates a new and deeply challenging security landscape. A traditional cyberattack targets data or system functions, for example, through jamming or data theft. A kinetic attack involves the application of physical force, such as an anti-satellite missile.

Now, consider a scenario where an OSAM servicer’s command and control system is compromised via a cyberattack. The attacker could then command the servicer to physically grapple and damage another satellite. Is this a cyber or a kinetic attack? It is fundamentally both. Similarly, an autonomous swarm could be commanded to surround a target satellite and use its low-power thrusters to subtly alter its orbit over time, a physical effect achieved through coordinated cyber commands.

This blurring of lines has profound consequences for international law, policy, and military rules of engagement. The foundational legal frameworks for space, including the Outer Space Treaty, were not designed to address such hybrid threats. This leaves a host of critical, unresolved questions that must become priorities for legal and policy research. For example, if a commercial OSAM vehicle from nation A, servicing a satellite from nation B, is hacked by a non-state actor in nation C and subsequently damages a satellite belonging to nation D, who is liable? The current international liability and responsibility frameworks are inadequate to address such a complex, multi-party scenario. Furthermore, at what point does a malicious cyber operation against a physically capable space asset cross the threshold to be considered a “use of force” under international law? The ambiguity of these hybrid threats creates a dangerously high risk of miscalculation and escalation, where a seemingly reversible cyber intrusion could provoke an irreversible physical conflict.

7.5 Policy, Governance, and Standardization

Technological solutions, no matter how advanced, are insufficient to secure the space domain in isolation. They must be developed and deployed within a robust and coherent framework of international policy, clear governance structures, and universally adopted standards. The global, interconnected, and interdependent nature of space operations means that a vulnerability in one nation’s system can pose a direct threat to all others. Establishing this framework is a critical prerequisite for a stable and secure future in space.

A fundamental challenge is that existing international space law, most notably the Outer Space Treaty of 1967, was drafted decades before the digital age and therefore does not explicitly address cybersecurity. This has created a significant legal and policy vacuum regarding critical issues like attribution for cyberattacks, liability for damages, and acceptable norms of behavior for cyber activities in space. Future work must focus on fostering international dialogue, for example through the United Nations Committee on the Peaceful Uses of Outer Space (COPUOS), to establish clear norms of responsible state behavior. This includes defining what constitutes a hostile act in the cyber domain and establishing clear channels and protocols for communication and de-escalation to prevent miscalculation.

Addressing the ambiguity of liability for cyberattacks is another paramount task. This will require a concerted effort to adapt principles from existing treaties, such as the state responsibility and absolute liability concepts from the Outer Space Treaty and the Liability Convention, to the unique context of the cyber domain. While legally and technically complex, establishing clear accountability is essential to deter malicious activity.

Alongside policy development, the promotion of universal technical standards is crucial for interoperability and baseline security. Bodies like the Consultative Committee for Space Data Systems (CCSDS), also known as ISO Technical Committee 20 Subcommittee 13, play a vital role in this process and should be supported in their efforts to develop and promulgate standards for secure command and control, authenticated data formats, and secure inter-satellite communication protocols. This work must be complemented by the harmonization of national-level regulations, such as the EU's proposed space cybersecurity laws and the principles outlined in U.S. Space Policy Directive-5, to create a consistent global security baseline and avoid a fragmented and inefficient patchwork of differing compliance requirements.

Finally, governance must extend to the entire lifecycle of a space system, with a particular focus on the supply chain. Policies must be implemented that mandate robust Cybersecurity Supply Chain Risk Management (C-SCRM) practices for all hardware and software components intended for space systems. Leveraging established frameworks like the NIST Cybersecurity Framework can provide a structured approach to identifying, assessing, and mitigating these risks from a system's inception, ensuring that security is built in, not bolted on.

7.6 Cyber Security Testbed

Recent anomalies from ground-side credential theft to on-orbit bus resets show that security faults in space missions seldom respect organisational or subsystem boundaries. Currently, the space security community lacks a realistic, easily accessible testbed. A scientifically sound *testbed* must therefore function as more than an engineering sandbox: it must be a research instrument that allows hypotheses about security, safety, and resilience to be stated precisely, evaluated systematically, and reproduced independently. The following six design dimensions structure this instrument and crucially explain *why* each is indispensable for scholarly work.

- Segment-complete, abstraction-aware modelling. Attack chains traverse the ground, link, and space segments, and omitting any segment hides entire classes of causality. Therefore, we demand explicit coverage of all three segments even when the RF link is abstracted precisely to keep cross-segment effects observable. Formally recording each segment and its interfaces means that researchers can vary fidelity locally (e.g., replace a hardware ground station with a stochastic delay channel) without invalidating global semantics.
- Graduated fidelity architecture. Simulation is fast and cheap, yet certain timing or radiation effects only surface when real avionics boards are in the loop. We suggest a sliding scale of fidelity that formalises this continuum depending on the research use case, e.g., incident response vs. attack forensics. By binding every experiment to a declarative manifest, a result obtained in a low-fidelity simulation can later be replayed.
- Executable realism with unmodified binaries. Many spacecraft exploits hinge on low-level behaviour (e.g., bus arbitration, watchdog timeouts) that disappears when flight software is re-linked for a laboratory harness. Therefore, it is essential that low-level access (binary) runs within the testbed. Embedding cycle-accurate processor models, flatsats, or physical boards creates an executable ground-truth layer against which analytical models can be calibrated.
- Formal scripting language for faults and threats. The testbed should use one clear language that can describe both random hardware faults and deliberate cyberattacks. By automating techniques from frameworks such as SPARTA, ESA SHIELD, and MITRE

ATT&CK, and mixing them with classic fault injections, we can measure exactly how much of the threat and fault space we have tested. Each scripted event is tagged and traced to its effect, letting us run solid statistics on how well proposed defences perform.

- Data-first instrumentation and stewardship. Empirical progress depends on transparent, multi-layer telemetry. We insist on capturing three data classes *metadata*, *nominal*, and *attack* traces for every experiment. At the same time, we suggest granular traffic-flow visibility for forensics and recovery research. Embedding synchronised probes at computation, communication, and energy layers satisfies these requirements and produces curated corpora for future machine-learning studies that currently lack representative data.
- Openness, sustainability, and community extensibility. Proprietary solutions fragment evidence and impede replication. Therefore, for the testbed, we advocate for open standards (CCSDS, ECSS) and a shared registry of benchmark scenarios. A plug-in architecture allows new sensor models, cryptographic stacks, or threat patterns to be added with negligible re-engineering effort, thus ensuring that the testbed evolves alongside mission technology.

8 Recommendations for Future Research and Development

The seminar established that space cybersecurity challenges differ in kind from terrestrial ones, and the working groups on threat preparation, detection, protection, and response surfaced actionable gaps. This section translates those findings into a sequence of interdependent pillars that government agencies and institutions, industry, and academia can use to build a cohesive and cumulative research ecosystem.

Pillar 1: Unified Research Agenda

Progress begins with a shared, public agenda that ties operational pain points to research tasks and evaluation criteria. Space agencies, operators, industry, and academic partners should co-author and annually refresh this agenda, explicitly including long-horizon topics such as secure key distribution for deep space and DTN, physics-informed anomaly detection, provenance-preserving fault tolerance, secure updateability, and verifiable assurance for autonomous systems so foundational science stays aligned with mission needs. This is fundamental for multiple reasons:

- Institutional agencies and industry are dependent on the availability of lower TRL research in order to execute higher TRL prototype development, production, and operationalization.
- Academia is dependent on use case scenarios and long-term visions of the institutional players and industry in order to be able to select relevant and impactful research topics

The unified research agenda can capture these dependencies and better link the various actors and their needs.

Pillar 2: Access to High-Fidelity Artifacts and Platforms

Empirical and reproducible research is contingent upon access to realistic *artifacts*, spanning not only telemetry data but also low-level system components. Government and commercial operators should prioritize the creation and dissemination of flight-like datasets. This should include artifacts from missions that are post-mission or have been deorbited, as the operational risk is eliminated and data can be re-evaluated for release under appropriate agreements. Such datasets should include releasable or anonymized telemetry logs, telecommands, internal satellite bus traffic (e.g., CAN bus, SpaceWire), and firmware images for key subsystems. Where proprietary constraints permit, access to redacted source code offers the highest level of ground truth for formal analysis. When direct release of these artifacts is not feasible, they should be used to curate high-fidelity synthetic corpora anchored to real mission parameters or be made available for analysis within secure data enclaves. Furthermore, a structured framework should be established to grant researchers hands-on access to realistic hardware. This includes creating a repository for retired Engineering Qualification Models (EQMs) from past missions. Additionally, dedicating time to security experiments on operational research platforms, akin to the OPS-SAT [2] model, provides invaluable data on real-world systems. A more ambitious step would be for operators to provide sanctioned research access to in-orbit spacecraft after their primary mission life has concluded. In return, the academic community must commit to the rigorous use of these scarce resources, including documenting dataset limitations and releasing open-source models and data generators to ensure results are comparable and verifiable. At the same time, the research community, supported by space agencies and industry, should work toward developing a space cybersecurity testbed that offers dynamic fidelity to accommodate diverse research experiments.

Pillar 3: Aligned Roles and Incentives.

A sustainable research ecosystem requires a collaborative framework that aligns the distinct roles and incentives of government, industry, and academia. In this tripartite model, government agencies and commercial operators provide the essential context by defining mission constraints, furnishing operational data, and brokering access to hardware. Industry partners serve as the crucial conduit for technology transition, contributing commercial-grade tools, standardized test vectors, and viable pathways to productization. The academic community provides the scientific foundation, delivering novel methods, open-source benchmarks, and the in-depth analysis required to address long-term challenges.

To be effective, collaborative agreements must be structured to recognize the different time horizons inherent to each sector. Projects should be designed to yield both near-term, tangible deliverables (such as software tools and test cases) and to support long-term, foundational research (such as the development of formal methods, principles for autonomy safety, and strategies for PQC migration). This model is designed to be mutually beneficial, creating a virtuous cycle of innovation and capability. Academia benefits from access to relevant problems and data, resulting in peer-reviewed publications and a highly skilled workforce. In return, government and industry gain access to independently validated technologies, robust security evaluations, and ultimately, a reduction in the risk and cost of integrating new security solutions into operational missions.

Pillar 4: Fit-for-Purpose Funding and Collaboration Models

A persistent gap between academic innovation and operational reality is the absence of sustained funding instruments dedicated to low-TRL (TRL 1–3) research. To bridge this, major national and international research programs, such as Horizon Europe and the US National Science Foundation (NSF), alongside other agencies, must establish dedicated, multi-year funding thrusts for space cybersecurity. These programs should be structured as competitive calls for academic-led projects, ensuring publishable results by default and producing deliverables that strengthen the entire research pipeline, including open benchmarks, reference implementations, and curated datasets.

A highly effective structure for these funded projects involves joint research programs, such as industry or agency co-funded PhD positions and fellowships, which embed academic researchers directly within operational environments under pre-negotiated intellectual property agreements. They should be tied to the unified research agenda. This model directly couples funding with commitments for access to data, experimenter time on spacecraft, and controlled use of Engineering Qualification Models (EQMs). By allowing researchers to work with sensitive hardware and data in situ, this approach solves the critical access problem, creating a virtuous cycle: academia gains invaluable access to real-world challenges, while agencies and industry de-risk new technologies and build a direct pipeline to specialized talent.

Pillar 5: Enforceable Standards and Governance.

Technological advances must be codified into enforceable standards and supported by clear governance to ensure a consistent and high security baseline across the space ecosystem. A critical starting point is through procurement and acquisition policy, which should mandate security-by-design principles from inception. Foundational requirements for new systems must include a hardware-rooted secure boot process, cryptographic agility with a clear migration path to Post-Quantum Cryptography (PQC), verifiable software update mechanisms, and policies for post-incident data retention. Furthermore, mandating minimal, interoperable logging and attestation schemas for both command links and internal buses is essential for future incident response and analysis.

Beyond individual systems, community-wide security depends on robust information sharing. This necessitates the creation of a space-focused threat intelligence exchange, building on existing standards such as STIX/TAXII but extending them with space-specific observables and Space Situational Awareness (SSA) context. To encourage proactive defense, clear "safe-harbor" policies for vulnerability disclosure should be established, providing legal protection for good-faith security researchers. Finally, as missions increasingly involve multiple commercial and international partners, unambiguous liability frameworks are required to make collaboration practical by assigning responsibility in the event of a security incident.

Finally, in particular, for solar system Internet scenarios, communication security solutions should be standardized through the Consultative Committee for Space Data Systems (CCSDS) to ensure maximum impact and interoperability.

9 Conclusion

This Dagstuhl Seminar convened 40 leading experts from academia, industry, and space agencies, many of whom possess significant experience securing terrestrial cyber-physical systems such as industrial control systems and autonomous vehicles. This diverse group reached a clear consensus: space cybersecurity is not an incremental extension of terrestrial challenges but a qualitatively distinct discipline. The unique interplay of a deceptive physical environment that creates attribution ambiguity, the operational necessity of high-stakes autonomy under extreme latency, and a uniquely asymmetric and expanding attack surface forge a security paradigm that demands a fundamental shift in our approach. To address these foundational challenges, the seminar's four working groups translated this diagnosis into concrete priorities.

Moving forward, progress cannot be achieved in isolated silos. The path to secure and resilient space systems is not paved by technological solutions alone but is built upon a strategic foundation of collaboration and shared resources. The recommendations outlined in this report, from establishing a unified research agenda and providing access to high-fidelity artifacts to aligning stakeholder incentives and creating fit-for-purpose funding or designing a testbed dedicated to space cybersecurity research, are not independent objectives but an interdependent roadmap. The central message of this seminar is a call to action: for space agencies, industry, and academia to collaboratively build the open, reproducible, and cumulative research ecosystem required to safeguard our critical infrastructure in orbit and beyond.

References

1 Nicolò Boschetti, Nathaniel G. Gordon, and Gregory Falco. Space cybersecurity lessons learned from the viasat cyberattack. In *ASCEND 2022*. American Institute of Aeronautics and Astronautics (AIAA), 2022.

2 David Evans and Mario Merri. Ops-sat: A esa nanosatellite for accelerating innovation in satellite control. In *SpaceOps 2014 Conference*, page 1702, 2014.

3 Courtney Fleming, Mark Reith, and Wayne Henry. Securing commercial satellites for military operations: A cybersecurity supply chain framework. In *Proceedings of ICCWS 2023: The 18th International Conference on Cyber Warfare and Security*, pages 85–92. Academic Conferences and Publishing Limited, 2023.

4 Max Planck Institute for Extraterrestrial Physics. ROSAT – the end of an exceptional satellite – mpe.mpg.de. `https://www.mpe.mpg.de/229897/News_20111114`. [Accessed 27-05-2025].

5 Loren Grush. A Russian satellite seems to be tailing a US spy satellite in Earth orbit – theverge.com. `https://www.theverge.com/2020/1/31/21117224/russian-satellite-us-spy-kosmos-2542-45-inspection-orbit-tracking`. [Accessed 28-05-2025].

6 Ralph Langner. Stuxnet: Dissecting a cyberwarfare weapon. *IEEE security & privacy*, 9(3):49–51, 2011.

7 Banks Lin, Wayne Henry, and Richard Dill. Defending small satellites from malicious cybersecurity threats. In *International Conference on Cyber Warfare and Security*, volume 17, pages 479–488, 2022.

8 Syed Shahzad, Keith Joiner, Li Qiao, Felicity Deane, and Jo Plested. Cyber resilience limitations in space systems design process: Insights from space designers. *Systems*, 12(10):434, 2024.

Participants

- Ali Abbasi
  CISPA – Saarbrücken, DE
- Steven Arzt
  Fraunhofer SIT – Darmstadt, DE
- Brandon Bailey
  The Aerospace Corp. – Los Angeles, US
- Lars Baumgärtner
  ESA / ESOC – Darmstadt, DE
- Nesrine Benchoubane
  Polytechnique Montréal, CA
- Simon Birnbach
  University of Oxford, GB
- Antonio Carlo
  Tallinn University of Technology, EE
- José Manuel Diez López
  TU Berlin, DE
- Knut Eckstein
  ESA / ESTEC – Noordwijk, NL
- Gregory J. Falco
  Cornell University – Ithaca, US
- Daniel Fischer
  ESA / ESOC – Darmstadt, DE
- Kevin Gilbert
  NASA – Greenbelt, US
- Florian Göhler
  BSI – Bonn, DE
- Arne Grenzebach
  OHB System – Bremen, DE
- Gürkan Gür
  ZHAW – Winterthur, CH
- Jessie Hamill-Stewart
  University of Bristol, GB
- Wayne “Chris” Henry
  Air Force Inst. of Technology – Wright-Patterson, US
- Eric Jedermann
  RPTU – Kaiserslautern, DE
- Samuel Jero
  MIT Lincoln Laboratory – Lexington, US
- Gunes Karabulut Kurt
  Polytechnique Montréal, CA
- Syed Ibrahim Khandker
  New York University – Abu Dhabi, AE
- Vincent Lenders
  armasuisse – Thun, CH
- Efrén López Morales
  Texas A&M University – Corpus Christi, US
- Mark Manulis
  Universität der Bundeswehr – München, DE
- Carsten Maple
  University of Warwick, GB & Alan Turing Institute – London, GB
- Ulysse Planta
  CISPA – Saarbrücken, DE
- Aanjhan Ranganathan
  Northeastern University – Boston, US
- Markus Rückert
  ESA / ESOC – Darmstadt, DE
- Peter Y. A. Ryan
  University of Luxembourg – Esch-sur-Alzette, LU
- Harshad Sathaye
  ETH Zürich, CH
- Stephen Schwab
  USC/ISI – Arlington, US
- Mridula Singh
  CISPA – Saarbrücken, DE
- Jill Slay
  University of South Australia – Mawson Lakes, AU
- Joshua Smailes
  University of Oxford, GB
- Fiona Stone
  UK Space Agency – London, GB
- Martin Strohmeier
  armasuisse – Thun, CH
- Rosa Szurgot
  Embry-Riddle Aeronautical University – Prescott, US
- Mattias Wallén
  Swedish Space Corporation – Solna, SE
- Marcus Wallum
  ESA / ESOC – Darmstadt, DE
- Johannes Willbold
  Ruhr-Universität Bochum, DE